

A ONTOLOGIA NA CIÊNCIA DA INFORMAÇÃO

Marta Sianes Oliveira do Nascimento
Geovane Eugênio Oliveira
Geraldino Gonçalves Bastos
Fernanda Passini Moreno

Resumo

Este artigo apresenta o uso do termo ontologia, a partir de levantamento da frequência deste em artigos de periódicos indexados na Base LISA. Procurou-se identificar a evolução do volume de publicações em diversos idiomas, desde a primeira vez em que o termo foi utilizado, em 1977, até março de 2006; a relação do termo com outros temas e a identificação dos pesquisadores com maior contribuição acadêmica na área. Os dados encontrados sugerem que deve ser realizado um estudo qualitativo, como a análise de conteúdo dos resumos ou textos completos para estabelecer uma análise mais aprofundada do escopo em que a ontologia se vê inserida hoje no campo da Ciência da Informação.

Palavras chave:

Ontologia; Ciência da Informação; Produção científica; Bibliometria

THE ONTOLOGY IN INFORMATION SCIENCE

Abstract

This paper analyzes the use of the term ontology based on the frequency it is used on articles of indexed journals at LISA database. The objective was to identify the increasing of publications in several languages, since the first time the term was used, in 1977, until March 2006. It also tries to identify the relation of ontology with other majors and the researchers that most contributed with this academic area. The founded data suggest that it must be developed a qualitative study, like abstracts or full texts contents' analysis to establish a deeper scope analysis where Ontology is inserted today at Information Science area.

Key-words:

Ontology; Information Science; Bibliometrics; Scientific Production

1 INTRODUÇÃO

O termo ontologia vem sendo discutido e aplicado em diversas áreas do conhecimento científico. Originado na filosofia, significa: “categoria, que pode ser usada para classificar alguma coisa” (BAX, 2003, p.7). Para HOUAISS (2006), “parte da filosofia [aristotélica] que tem por objeto o estudo das propriedades mais gerais do ser”.

Atualmente, o termo adquire nova roupagem e:

São encontradas na literatura diversas definições para as ontologias, diversos tipos, propostas para aplicação em diferentes áreas de conhecimento e propostas para a construção de ontologias (metodologias, ferramentas e linguagens). (ALMEIDA; BAX, 2003, p.7)

Conforme apresentou Gruber (1993), ontologia seria uma especificação explícita de uma conceituação, ou seja, a definição explícita de conceitos e suas relações, propriedades e restrições expressas formalmente. Segundo o autor, a ontologia permite especificar uma visão abstrata e simplificada de um assunto que se deseja representar. Depreende-se daí a grande variedade de aplicações para a ontologia, como pode ser verificado no QUADRO 1, a seguir.

QUADRO 1
Aplicações da ontologia

ÁREA DE APLICAÇÃO	DEFINIÇÃO
Inteligência Artificial	GUARINO (1997) define ontologia como uma caracterização axiomática do significado do vocabulário lógico
Sistemas de Informação	É definida como um conjunto de conceitos e termos que podem ser usados para descrever alguma área do conhecimento ou construir uma representação para o conhecimento (SWATOUT, 1999)
Linguagem e Cognição	A ontologia refere-se a tudo que existe no mundo composto por objetos, mudanças e relações entre eles. Pode ser baseada no mundo, na mente/intelecto, na cultura ou na linguagem (DAHLGREN, 1995).
Banco de Dados	Conhecimento genérico que pode ser reusado em aplicações de tipos diferentes (MEERSMAN, 2002)

Fonte: VITORINO; BRÄSCHER (2006)

Nicola Guarino (1998) acrescenta que o uso das ontologias habilita o desenvolvedor a praticar o alto nível de reuso que é problema constante em engenharia de software, permitindo ainda compartilhar aplicações de domínio do conhecimento, singularmente utilizadas no vocabulário aplicado às plataformas de software heterogêneas.

Conforme a Figura 1, abaixo, Guarino (1998) classifica os tipos de ontologias conforme seu nível de dependência sobre uma tarefa particular ou um ponto de vista. Desta forma, faz a distinção entre alto nível, domínio, aplicação e tarefas ontológicas.

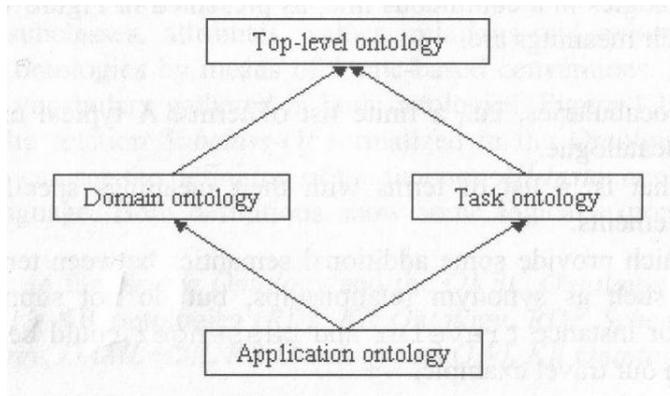


FIGURA 1 – Categorização

Fonte: GUARINO (1998)

A partir desta visão de diversas áreas do conhecimento que se valem do conceito de ontologia, impôs-se o seguinte questionamento: como a ontologia é estudada na Ciência da Informação? Quais autores têm produção científica sobre o tema? Em quais periódicos podem ser encontrados? O que vem sendo discutido sobre ontologia na área de CI?

O presente trabalho buscou responder a essas questões por meio do levantamento bibliográfico em bases de dados da área da Ciência da Informação (CI), com o intuito de perceber:

- usos do termo ontologia;
- evolução das publicações sobre o tema ao longo dos anos;
- frequência das palavras-chave associadas aos artigos científicos que tivessem como objeto estudos relacionados a ontologia.

O estudo objetivou, ainda:

- obter dados quantitativos acerca da produção dos autores;
- obter dados quantitativos sobre a língua e país onde originalmente foram publicados os artigos;
- levantar dados sobre quais periódicos se destacam na publicação de artigos relacionados ao tema e em quais anos ou décadas se concentram estes artigos.

Aliada à análise quantitativa, procedeu-se à análise da produção, considerando autores, palavras-chave, ano, idioma e periódico das publicações visando complementar com informações qualitativas os dados obtidos.

O tipo de pesquisa, os procedimentos metodológicos adotados, bem como a descrição das etapas e das fontes da pesquisa, são apresentados a seguir.

2. METODOLOGIA

Para o levantamento bibliográfico, optou-se por pesquisar, principalmente, em bases de dados referenciais e de resumos que indexam e organizam periódicos e textos científicos na área de CI.

De acordo com a tipologia de McGrath, (apud MACIAS-CHAPULA,1998), a pesquisa realizada poderia ser enquadrada como informétrica, considerando que as fontes utilizadas foram palavras, documentos e bases de dados, objetos de estudo desta ciência.

No entanto, o delineamento da pesquisa indicou a necessidade de uso de uma lei bibliométrica, a Lei de Zipf, que consiste em medir a frequência das palavras em textos, gerando uma lista ordenada de termos de determinada disciplina ou assunto.

Caracterizou-se o estudo, portanto, como bibliométrico, pois como afirma Macias-Chapula (1998, p. 134) "Bibliometria é o estudo dos aspectos quantitativos da produção, disseminação e uso da informação registrada.

2.1 Levantamento Bibliográfico

Os levantamentos nas bases de dados foram realizados no período de maio-junho de 2006. Entre as diversas bases da área de Ciência da Informação, optou-se por limitar a pesquisa em duas bases dependentes de convênio para acesso (bases de dados comerciais) e três bases ou repositórios de acesso livre. As bases selecionadas deveriam fornecer o campo de resumo dos artigos indexados.

Na primeira fase, foram pesquisadas diferentes bases de dados com o termo 'ontology'. Tal fato levou a uma alta revocação¹, posto não terem sido adotados filtros de nenhuma natureza nas opções de pesquisa, isto é, os campos pesquisáveis como título, resumo e palavra-chave foram considerados válidos para a pesquisa.

A próxima subseção traz a descrição das fontes utilizadas.

2.1.1 Bases de dados

LISA - Library and Information Abstracts

<http://www.csa.com/factsheets/lisa-set-c.php>

Sediada na Inglaterra, acessível por meio da Internet (*login* através de instituições consorciadas a CAPES), disponível em CD-ROM ou através de contato com a Biblioteca do IBICT. Trata-se de uma das bases de dados referencial mais acessada na área de CI, com cobertura internacional. Atualmente, indexa mais de 440 periódicos publicados em 68 países. O período disponível on-line é de 1969 até os dias atuais (IBICT, 2006)

DoIS - Documents in Information Science

<http://wotan.liu.edu/doi/>

Serviço para busca e *download* de publicações em CI. A base de dados contém 13963 artigos e 4493 *proceedings* de conferências, sendo que 13041 estão disponíveis para *download*. A DoIS é resultado de um esforço voluntário de criação de uma bibliografia de acesso livre de textos especializados em CI.

E-LIS – E-prints in Library and Information Science

<http://eprints.rclis.org/>

Repositório de acesso livre de documentos científicos e técnicos, publicados ou não publicados sobre Biblioteconia, CI, Tecnologia e área correlatas, resultado do esforço voluntário. Base não-comercial, possui atualmente 3804 documentos.

¹ " Habilidade de recuperar informação relevante para um assunto específico, quando uma solicitação sobre esse assunto é feita ao sistema. (LANCASTER, 1977) " (GLOSSÁRIO GERAL DE CIÊNCIA DA INFORMAÇÃO, 2006)

ISTA - Information Science & Technology Abstracts

<http://www.lsuhs.edu/no/library/ss&d/data/ista.html>

Indexa mais de 450 periódicos na área de Ciência da Informação. Inclui livros, relatórios de pesquisa, congressos e conferências, patentes e outros materiais (IBICT, 2006).

Google Acadêmico

<http://scholar.google.com.br/intl/pt-BR/scholar/about.html>

A ferramenta indexa artigos revisados por especialistas (*peer-reviewed*), teses, livros, resumos e artigos de editoras acadêmicas, organizações profissionais, bibliotecas de pré-publicações, universidades e outras entidades acadêmicas. A tecnologia de classificação do Google leva em conta o texto integral de cada artigo, o autor, a publicação em que o artigo saiu e a frequência com que foi citado em outras publicações acadêmicas. Não há dados sobre os números de artigos ou documentos indexados.

2.2 Procedimentos para coleta e organização dos dados

Para as pesquisas efetuadas nas bases LISA e ISTA, dependentes de convênio entre instituições para o acesso, foi contatada a biblioteca do IBICT e solicitada a busca em ambas as bases com o termo 'ontology', sem a utilização de filtros como ano, por exemplo. Como afirmado anteriormente, buscou-se perceber a ocorrência da palavra nos campos: título, resumo e palavra-chave.

As buscas em repositórios livres e temáticos de Ciência da Informação, E-LIS e DoIS, foram realizadas por meio de acesso à internet, valendo-se da mesma palavra-chave, e igualmente sem restrições ou filtros. Nas buscas realizadas na ferramenta acadêmica do Google foram feitas duas combinações: 'ontology' e 'ontology' and 'information science', na segunda pesquisa, por não se tratar de base de dados restrita à área de CI.

O número de registros encontrados em cada base de dados pode ser verificado no quadro abaixo:

QUADRO 2

Número de registros encontrados nas bases de dados pesquisadas

BASE DE DADOS	Nº DE REGISTROS (1)	Nº DE REGISTROS (2)
DoIs	12	1
E-Lis	08	41
ISTA	148	-
LISA	288	-

- (1) ocorrências expressas no título, no resumo e/ou na palavra-chave
- (2) ocorrências expressas apenas no texto integral da publicação

Na busca realizada no Google Acadêmico que considerou o termo ‘ontology’, o resultado totalizou 325.000 registros. Na segunda busca, associando o termo a ‘information science’, o resultado apresentado totalizou 6.210 registros.

Após este levantamento inicial, obteve-se uma visão geral de ocorrência do termo. Como foi apresentada uma alta recuperação no conjunto das bases pesquisadas, alguns recortes se impuseram: a análise detalhada dos repositórios livres foi descartada, acreditando que, potencialmente, poderiam duplicar os resultados. O mesmo se deu em relação à ferramenta de pesquisa do Google, pois apontaria, provavelmente, resultados que já haviam ocorrido em outras bases. A base ISTA foi descartada para o estudo por conter materiais irrelevantes para o tema em estudo.

Dada a alta revocação encontrada, procedeu-se à segunda etapa: seleção da base de dados que seria analisada em profundidade. Optou-se por restringir a análise à base de dados LISA, que indexa as principais revistas da área e, refletindo, portanto a evolução da pesquisa em CI.

Os dados obtidos na base LISA foram organizados em planilha eletrônica, de acordo com as seguintes categorias: ano, autor, periódico, país, título do artigo e palavra-chave e produzidas informações estatísticas sobre o espectro de pesquisa.

É importante destacar que a base LISA não traz a informação do país da publicação, mas o idioma em que foi publicado o artigo. Dessa forma, na organização dos dados foi considerado o idioma da publicação e não o país do autor. Essa dificuldade apresenta uma limitação do estudo, pois é consenso que autores de várias nacionalidades publicam preferencialmente no idioma inglês.

3 Análise e Apresentação dos Resultados

A base LISA recuperou 288 registros com a busca do termo 'ontology'. No entanto, considerando algumas inconsistências, relatadas a seguir, o estudo utilizou 280 artigos para apresentação e análise dos resultados.

3.1 Considerações Acerca dos Resultados Encontrados

3.1.1 O mesmo artigo produzido em 1992, em inglês, por Dahlberg, I., encontra-se publicado em dois periódicos: *Information Services and Use*, v. 12, n.1, p.65-71, e o *International Classification*, v. 19, n.2, p.65-71. Por considerarmos essa repetição uma inconsistência, apenas um registro foi considerado no cômputo dos resultados.

3.1.2 Houve a ocorrência de dois registros iguais indexados em 2005, de autoria de Denda, K., publicado no periódico *Library Resource & Technical Services*, v. 49, n.4, out. 2005. Um dos registros foi eliminado da planilha para a apuração dos resultados.

3.1.3 Três artigos foram identificados na base sem informação sobre o idioma, conforme o quadro a seguir. Os três foram desconsiderados nos resultados analisados.

QUADRO 3

Artigos sem informação sobre o idioma de origem

Ano	Autor	Título do Periódico	Título do Artigo
2004	Gómez-Perez, Asuncion; Lozano-Tello, Adolfo	Journal of Databases e Management, v.15, n.2, p.1-18, Apr/July, 2004	Ontometric: a method choose the appropriate ontology
2004	Maillot, N.; Thonnat, M.; Boucher, A.	Machine Vision and Applications, v.16, Part 1, p.33-40, 2004	Towards ontology-based cognitive vision
2005	Geivas, P.; Diaz-Agudo, B.; Peinaldo, F.; Hervas, R.	Knowledge Bases Systems, v.18, n.45, p.235-242, Aug. 2005	Story plot generation based on CBR

3.1.4 Foi identificado um artigo na base sem informação sobre o autor, conforme o QUADRO 4 a seguir. O artigo não foi considerado nos resultados

QUADRO 4

Artigos sem informação sobre o autor

Ano	Título do Periódico	Título do Artigo	Palavras-chave
2003	Information Management Report, Nov. 2003, p.7-8	W3C issues OWL recommendations	World Wide Web; Ontologies; Standards; OWL

3.1.5 Três artigos foram identificados na base sem informação sobre a palavra-chave e não foram considerados nos resultados. O artigo de Maillot, já havia sido eliminado por não constar o idioma da publicação, conforme o quadro anterior.

QUADRO 5

Artigos sem informação sobre palavra-chave

Ano	Autor	Título do Periódico	Título do Artigo
2004	Edgington, T.; Choi, B.; Henson, K.; Raghu, T.S.; Vinze, A.	Communications - ACM, v.47, n.11, p.85- 91, 2004	Adapting ontology to facilitate knowledge sharing
2004	Kwon, J; Jeong, D.; Lee, L.S., Baik, D.K.	International Journal of Software Engineering and Knowledge Engineering, v.14, Part 5, p.519-542, 2004	Intelligent semantic concept mapping for semantic query rewriting/optimization in ontology- based information integration
2005	Maillot, N.; Thonnat, M.; Boucher, A.	Machine Vision and Applications, v.16, Part 1, p.33-40, 2004	Towards ontology-based cognitive vision

Dessa forma, dos 288 registros iniciais obtidos, serão apresentados e analisados, nas próximas seções, 280 artigos, nas categorias definidas para o estudo: ano de publicação, idioma, autor, periódico e palavra-chave.

3.2 Ano da publicação

O período de estudo está limitado entre os anos de 1977 e 2006. Até junho de 2006, foram indexados na base estudada artigos até o mês de março.

O primeiro artigo encontrado na base LISA que aborda o tema ontologia, intitulado “*The developing cognitive viewpoint in information science*”, de autoria de Bertram Brookes, data de 1977. O termo ocorre no resumo, e sua leitura indicou que o conceito de ontologia pode ser entendido como uma taxonomia, pois é associado à categorização Popperiana dos três mundos.

O segundo registro em que o termo ocorre, “*The foundation of information science, part I: philosophical aspects*” é um artigo de 1980, igualmente de autoria de Bertram Brookes, novamente relacionado à teoria Popperiana, como pode ser comprovado neste trecho do resumo: “*The Popperian ontology then has to be extended to admit the concept of information and its relation to subjective and objective knowledge.*”

Somente sete anos após esta publicação, em 1987, encontramos um artigo, em italiano, que utiliza o termo: “*Information Coding Classification. Un sistema di classificazione dai fondamenti ontologic*”, publicado no periódico italiano *Bibliotecario*, de março/junho de 1987, por Maria Teresa Biagetti. Da leitura do resumo depreende-se que aborda o esquema de classificação *Information Coding Classification* (ICC), de Ingetraut Dahlberg. O trecho, a seguir, traduzido do resumo, ilustra a ocorrência do termo: “Traça o desenvolvimento filosófico do ICC e a influência de Nicolai Hartmann no pensamento de Dahlberg a respeito da ontologia.” No QUADRO 6, são apresentadas as referências desses três artigos.

QUADRO 6

Três primeiros artigos sobre Ontologia indexados na base LISA

Ano	Autor	Periódico	Idioma	Título
1987	Biagetti, M. T.	Bibliotecário	Italian	Informations Coding Classification. A classification system with an ontological basis
1980	Brookes, B. C.	Journal of Information Science	English	The foundation of informations science, part 1: philosophical aspects
1977	Brookes, B.C.	Journal of Informatics	English	The developing cognitive viewpoint in information science

No Gráfico 1, a seguir, pode ser verificada a frequência de publicações de 1977 a março de 2006. Observa-se que, após 1987, somente em 1991 aparece outro artigo sobre o tema. A partir de 1991 houve publicações em todos os anos e apenas a partir de 2001, observa-se uma tendência crescente de publicações sobre ontologia.

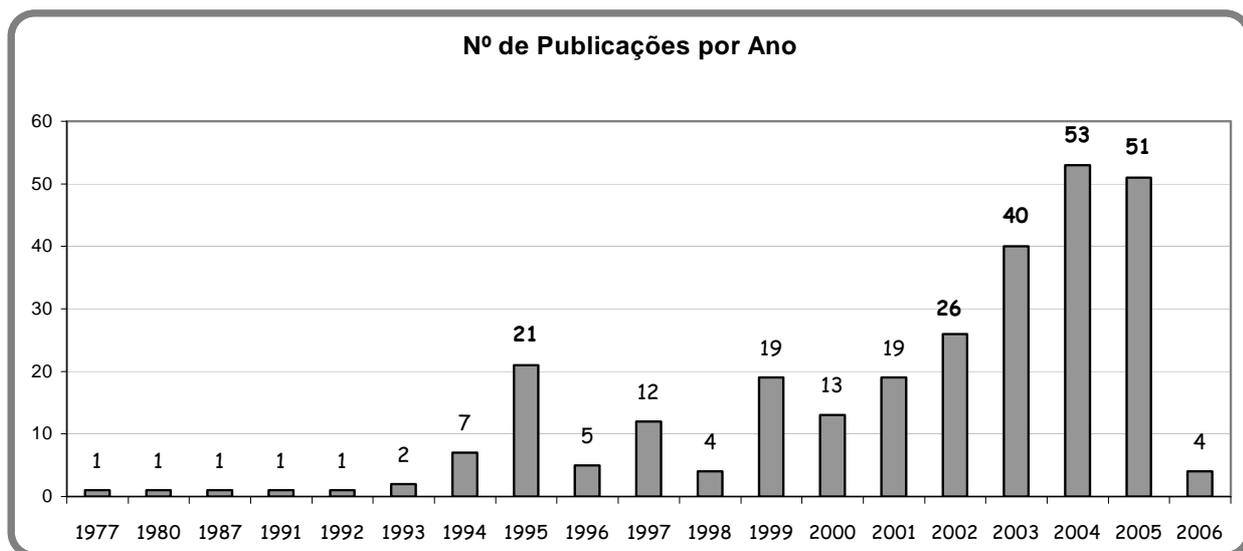


GRÁFICO 1 – Resultados por ano de publicação

*Os dados de 2006 referem-se aos artigos indexados na base LISA até março de 2006.

Conforme pode ser verificado no Gráfico acima, entre 1977 e 1994 – período de 17 anos –, 15 artigos, apenas, foram publicados sobre o tema.

Em 1995, foram publicados 21 artigos, todos no idioma inglês, sendo dezoito no periódico “*International Journal of Human-Computer Studies*”, edição de novembro/dezembro, dois no “*Artificial Intelligence in Medicine*”, edição de junho e um no “*Methods of Information in Medicine*”, em sua edição de março.

3.3 Idioma

Na análise da categoria idioma, foram encontradas publicações em inglês, russo, chinês, espanhol, alemão, italiano, húngaro, norueguês, japonês, português e africano.

De modo bastante significativo, o idioma inglês tem uma preponderância absoluta sobre os demais. Conforme detalhado no QUADRO 7, 82,52% das publicações encontradas na base LISA foram em inglês, contra pouco mais de 17% de publicações nos demais 10 idiomas encontrados.

Além disso, desde o primeiro ano de publicação, 1977, até hoje, excetuado os anos em que não houve publicações sobre o tema, há artigos publicados em inglês. Ressalta-se que em 1987, o único artigo publicado foi no idioma italiano.

Os resultados encontrados na categoria Idioma de publicação mostram a supremacia do idioma inglês na área em estudo, seguido pelos idiomas russo e chinês. No entanto, na análise feita em relação à categoria Autores, pode ser constatado que o idioma nativo dos autores não está refletido no idioma de publicação. Chineses, espanhóis, brasileiros e pesquisadores de todas as partes do mundo publicam em inglês.

QUADRO 7
Resultados por idioma de publicação

Idioma	Nº	%
English	231	82,21
Chinese	15	5,34
Russian	15	5,34
Spanish	5	1,78
German	4	1,42
Italian	3	1,07
Hungarian	2	0,71
Norwegian	2	0,71
Japanese	1	0,36
Portuguese	1	0,36

Idioma	Nº	%
Chinese	15	30,61
Russian	15	30,61
Spanish	5	10,20
German	4	8,16
Italian	3	6,12
Hungarian	2	4,08
Norwegian	2	4,08
Japanese	1	2,04
Portuguese	1	2,04
Afrikaans	1	2,04

Total exceto inglês: 49 publicações

No QUADRO 8, a seguir, pode ser observada a relação entre Idioma e o Ano de Publicação, considerando a primeira e a última publicação em cada um dos idiomas.

QUADRO 8
Relação Ano e idioma de Publicação

Idioma	Primeira Publicação	Última Publicação
English	1977	2006
Chinese	1996	2005
Russian	2000	2005
Spanish	1999	2000
German	2003	2005
Italian	1987	2005
Hungarian	2004	2004
Norwegian	2003	2003
Japanese	2003	2003
Portuguese	2001	2001
Afrikaans	1996	1996

Observa-se que as publicações nos idiomas africano, português, húngaro, norueguês, espanhol e japonês não apresentam continuidade nas publicações. Não se pode concluir, entretanto, que autores desses países interromperam sua produção científica. Pode-se, mais uma vez, constatar que no tema ‘ontologia’ na área de CI, o idioma de publicação reconhecido é o inglês.

Cabe mencionar que há uma preponderância de revistas cujo idioma de publicação é o inglês, o que pode justificar o predomínio desta língua. Somente uma análise do país de origem dos autores poderia refletir melhor o desenvolvimento do tema nas diversas línguas.

O único artigo indexado no idioma português – “Uma ontologia comum para a integração de bases de informações e conhecimento sobre ciência e tecnologia” – é de autoria de Roberto Carlos dos Santos Pacheco e Vinícius Medina Kern e foi publicado na revista *Ciência da Informação* (v.30 n.3) em 2001.

3.4 Autores

Na análise da categoria autor, procurou-se verificar os autores mais produtivos ao longo do período de estudo. O QUADRO 9 apresenta o total de publicações por autor, especificando o idioma de publicação e se o autor é apresentado como principal ou secundário no artigo indexado na base LISA. Estão representados apenas os autores que possuem pelo menos duas publicações sobre o tema em questão, como autor principal.

QUADRO 9
Produtividade dos autores

Autor	Idioma	Total	Princ.	Sec.	Autor	Idioma	Total	Princ.	Sec.
Artem'eva	russo	10	3	7	Valente	inglês	2	2	
Kleshchnev	russo	8	8		Stojanovic	inglês	2	2	
Noy	inglês	7	3	4	Qin, J	inglês	2	2	
Gomez-Perez	inglês	5	2	3	Knorz	alemão	2	2	
Poli	inglês	5	3	2	Heijst	inglês	2	2	
Staab	inglês	5	2	3	Handschuh	inglês	2	2	
Knyazeva	russo	5	2	3	Gruninger	inglês	2	2	
Guarino	inglês	5	3	2	Gangen	inglês	2	2	
Motta	inglês	4	2	2	Fu, L	chinês	2	2	
Kupnevich	russo	4		4	Eriksen	norueguês	2	2	
Dyng, Y	inglês	3	3		Colomb	inglês	2	2	
Maedche	inglês	3	2	1	Burstein	inglês	2	2	
Domingue	inglês	3	2	1	Brookes	inglês	2	2	
Dahlberg	inglês	2	2		Alani	inglês	2	2	
Rojas	espanhol	2	2						

Da visualização deste quadro constata-se que embora mais de 80% das publicações seja no idioma inglês (231 publicações), quando considerados individualmente, os dois autores com maior número de publicações são os russos Artem'eva (10) e Kleshchnev (8 artigos publicados como autor principal). Noy, autor que publica em inglês, aparece em terceiro lugar com 7 artigos indexados.

Foi percebida a existência de um grupo de autores que publicaram apenas em russo: Artem'eva, Kleshchnev, Knyazeva e Kupnevich. Esses autores foram responsáveis por 13 das 15 publicações em russo. Os outros dois pesquisadores russos - Paducheva e Uzileviskii - publicaram sozinhos. Artem'eva e Kleshchnev escreveram 7 artigos em parceria, Artem'eva, Knyazeva e Kupnevich três artigos e Knyazeva e Kupnevich dois artigos. Todos foram publicados no periódico russo *Nauchno - Tekhnicheskaya Informatsiya*.

Outras parcerias foram identificadas: Natalya Noy publicou um artigo com Nicola Guarino e Gomez-Perez em 1994 no *IEEE Intelligent Systems*; Nicola Guarino e R. Poli têm um artigo publicado, em 1995, no *International Journal of Human-Computer Studies*; S. Staab publicou com A. Maedche dois artigos, um em 2000 e outro em 2001 e Enrico Motta, publicou dois artigos com John Domingue, em 2000.

Y Ding publicou três artigos em inglês no *Journal of Information Science*, sendo dois artigos com Foo. Nenhum deles apresentou registro de publicação em chinês.

Além dos brasileiros R. C. dos Santos Pacheco e Vinícius M. Kern, que publicaram um artigo em português em 2001, Alexandra Moreira, Lídia Alvarenga e Alcione de Paiva Oliveira, autoras brasileiras, publicaram em inglês, o artigo "*Thesaurus and ontology: a study of the definitions found in the computer and information science literature, by means of an analytical-synthetic method*" no periódico "*Knowledge Organization*", em 2004.

O estudo sobre os autores que publicam sobre ontologia merece um aprofundamento, em especial na verificação de seu país de origem, parcerias entre autores de origens diferentes, vínculos com universidades e na análise dos temas de estudo.

3.5 Periódicos

No estudo, foram identificados 85 periódicos diferentes, sendo que 67 (78,82%) publicam na língua inglesa. Os resultados estão expressos no QUADRO abaixo.

QUADRO 10

Número de Periódicos por Idioma

Idioma	Nº	%
English	67	78,82
Chinese	3	3,53
Spanish	3	3,53
Italian	3	3,53
German	2	2,35
Norwegian	2	2,35
Russian	1	1,18
Hungarian	1	1,18
Japanese	1	1,18
Portuguese	1	1,18
Afrikaans	1	1,18

O único periódico de língua portuguesa indexado na base LISA, relacionado ao tema Ontologia, é a revista brasileira “Ciência da Informação”, com um artigo publicado em 2001.

Dado o grande número de periódicos, cabe destacar que, conforme pode ser observado no Quadro 11, a média de publicações em inglês está na faixa de 3,5 por periódico, enquanto no idioma chinês é 5 e no idioma russo é de 15.

QUADRO 11

Percentagem de Publicações por Periódico

Idioma	Total de Publicações	Nº de Periódicos	Publicação por Periódico
English	231	67	3,45
Chinese	15	3	5
Russian	15	1	15
Spanish	5	3	1,67
German	4	2	2
Italian	3	3	1
Hungarian	2	1	2
Norwegian	2	2	1
Japanese	1	1	1
Portuguese	1	1	1
Afrikaans	1	1	1
Total	280	85	3

O QUADRO 12, a seguir, mostra esses resultados mais significativos em relação ao número de publicações por periódico, considerando a percentagem em relação ao total de artigos considerados neste estudo.

Entre os principais periódicos analisados, destacam-se como mais produtivos e relevantes, dentro do critério de número total de publicações no periódico: 3 periódicos de idioma inglês: *International Journal of Human-Computer Studies* (50 artigos), *IEEE Intelligent Systems* (33 artigos) e *AI Magazine* (15 artigos), um periódico de idioma russo *Nauchno- Tekhnicheskaya Informatsiya* (15 artigos), um periódico que publica em chinês *Journal of the China Society for Scientific and Technical Information* (13 artigos), e outro de idioma inglês, *Artificial Intelligence in Medicine*, com 10 artigos publicados.

QUADRO 12

Periódicos com maior número de publicações

Título do Periódico	Língua	Nº	%
International Journal of Human-Computer Studies	English	50	17,79
IEEE Intelligent Systems	English	33	11,74
AI Magazine	English	15	5,34
Nauchno- Tekhnicheskaya Informatsiya	Russian	15	5,34
Journal of the China Society for Scientific and Technical Information	Chinese	13	4,63
Artificial Intelligence in Medicine	English	10	3,56

No QUADRO 13, estão listados todos os periódicos indexados na base LISA que publicam em idiomas diferentes do inglês, com o número total de publicações em cada um, no período e tema em estudo.

QUADRO 13

Periódicos que publicam em idiomas diferentes de Inglês

Periódico	Língua	Nº
South African Journal of Library and Information Science	Afrikaans	1
Bulletin of the Library Association of China	Chinese	1
Journal of Information, Communication, and Library Science	Chinese	1
Journal of the China Society for Scientific and Technical Information	Chinese	13
NFD Information: Wissenschaft und Praxis	German	1
Wissenschaft und Praxis	German	3
Tudományos és Muszaki Tajékoztató	Hungarian	2
Bibliotecario	Italian	1
Biblioteche Oggi	Italian	1
Bollettino AIB	Italian	1
Pharmaceutical Library Bulletin (Yakugaku Toshokan)	Japanese	1
DF Revy	Norwegian	1
Synopsis	Norwegian	1
Ciência da Informação	Portuguese	1
Nauchno- Tekhnicheskaya Informatsiya	Russian	15
Investigacion Bibliotecologica: Archivonomia, Bibliotecologia, e Informacion	Spanish	1
Revista Interamericana de Bibliotecologia	Spanish	2
Scire; 6 (1) Jan-Jun 2000	Spanish	2

É interessante destacar que o periódico chinês *Journal of The China Society for Scientific and Technical Information* e o russo *Nauchno-Tekhnicheskaya Informatsiya* publicaram 3 artigos cada em 2005. O periódico italiano *Biblioteche Oggi* e o periódico alemão

Wissenschaft und Praxis aparecem com uma publicação, cada um, em 2005. Os demais periódicos apresentados no QUADRO 13 não apresentam publicações em 2005 e, em 2006, só estão indexados periódicos que publicam no idioma inglês.

3.6 Palavras-Chave

Na categoria Palavra-chave, buscou-se identificar as mais frequentes nos artigos recuperados e, também, identificar as áreas às quais o termo “*ontology*” pode estar relacionado.

As palavras “*ontologies*” e “*ontology*” apareceram 128 vezes na categoria Palavra-chave da base LISA, representando aproximadamente 10% do total de 1252 palavras-chave da base estudada. Do universo de 616 palavras, apenas 26 representam 49,20% do total de palavras-chave, as quais se repetem 10 ou mais vezes.

A análise da categoria Palavra-chave revestiu-se de grande dificuldade, uma vez que o número de palavras-chave por artigo era bastante numeroso variando de três a dez palavras-chaves. Esta característica conduz à constatação da necessidade de aprofundar os estudos no sentido de realizar uma análise de conteúdo nos artigos, seja pelo *abstract* ou obtendo o texto integral, para identificar, com clareza, as áreas de conhecimento e áreas de aplicação relacionadas ao termo.

Conforme pode ser observado no QUADRO 14, os termos mais citados além de “*ontologies*” e “*ontology*” se referem a palavras relacionadas à rede mundial de computadores, aplicações em computador, inteligência artificial, busca, gestão e representação do conhecimento.

QUADRO 14

Incidência de palavras-chave

<i>Quadro Resumo, incidência de palavras-chave</i>		
Palavra-chave	Nº	% do total
10 ou mais incidências (Total)	616	49,20
Ontologies	74	5,91
Ontology	54	4,31
World Wide Web	47	3,75
Computer applications	46	3,67
Artificial intelligence	36	2,88
Knowledge management	31	2,48
Knowledge based systems	24	1,92
Searching	24	1,92
Knowledge representation	23	1,84
Semantic web	22	1,76
Expert systems	19	1,52
Internet	19	1,52
Medicine	17	1,36
User interface	17	1,36
Computerized information storage and retrieval	15	1,20
Human-computer interaction	15	1,20
Online information retrieval	15	1,20
Software	15	1,20
Theories	15	1,20
Knowledge	14	1,12
Models	14	1,12
Information technology	13	1,04
Metadata	13	1,04
Semantic relations	12	0,96
XML	12	0,96
Computer science	10	0,80

<i>Incidência de palavras-chave na base</i>			
Incidência	Nº	Produto	% Total
10 ou mais incidências		616	49,20
9	3	27	2,16
8	4	32	2,56
7	4	28	2,24
6	3	18	1,44
5	8	40	3,19
4	13	52	4,15
3	22	66	5,27
2	52	104	8,31
1	269	269	21,49
Total		1252	100,00

Com o objetivo de apresentar brevemente as principais áreas de estudo e de atividades relacionadas aos artigos recuperados, alguns termos que apresentaram grande incidência como palavras-chave, sozinhos ou associados a outros termos – *Knowledge*, *Computer e Information* – serão exibidos de forma destacada nos QUADROS a seguir.

QUADRO 15
Palavras-Chave com o termo *Knowledge*

Palavra-chave	Nº	%
Knowledge	105	37,37
knowledge management	31	11,0
knowledge based systems	24	8,5
knowledge representation	23	8,2
"knowledge"	14	5,0
knowledge acquisition	5	1,8
knowledge organization	5	1,8
knowledge discovery	2	0,7
classification of knowledge	1	0,4

QUADRO 16
Palavras-Chave com o termo *Computer*

Palavra-chave	Nº	%
Computer	101	35,9
Computer applications	46	16,4
computadorized information storage and retrieval	15	5,3
human-computer interaction	15	5,3
computer science	10	3,6
computers	6	2,1
computer assisted instruction	3	1,1
computer supported colaborative work	2	0,7
computer assisted software engineering	2	0,7
computer security	1	0,4
computer production	1	0,4

QUADRO 17
Palavras-Chave com o termo *Information*

Palavra-chave	Nº	%
Information	63	22,4
online information retrieval	15	5,3
information technology	13	4,6
information work	8	2,8
information science	7	2,5
information communication	5	1,8
information systems	4	1,4
library and information science	3	1,1
management information systems	2	0,7
"information"	1	0,4
access to information	1	0,4
business information	1	0,4
online information storage	1	0,4
ontology-based Legal Information Environment	1	0,4
scene of Crime Information System	1	0,4

Dentre os artigos que tratam de ontologia, o termo “*information*” sozinho ou associado a outros termos aparece como palavra-chave desde 1977. Em 1999, 2000 e 2001 surge a expressão “*online database*”, mas apenas em 2002 a expressão “*on line information retrieval*” aparece como palavra-chave. É interessante observar que, embora o termo “*knowledge*” esteja presente como palavra-chave desde 1977, somente em 1994 veio associado a outros termos como “*knowledge based systems*” e “*knowledge representation*” e apenas em 1995 apareceu a expressão “*knowledge management*” como palavra-chave, mantendo-se presente até 2006. As expressões “*artificial intelligence*” e “*computer applications*” apareceram pela primeira vez como palavras-chave em 1994. O primeiro artigo que apresenta o termo “*web*” como palavra-chave data de 1997, sendo que apenas em 2001 aparece a expressão “*web semantic*” como palavra-chave.

No QUADRO 18, são apresentadas algumas palavras-chave agrupadas em relação a alguns usos de Ontologias.

QUADRO 18

Palavras-chave associadas a alguns usos de Ontologias

Palavra-chave	Nº	%
retrieval	34	12,1
storage	17	6,0
user interface	17	6,0
analysis	15	5,3
reuse	7	2,5
integration	6	2,1
interoperability	4	1,4
intelligent agents	4	14,0

As percentagens dos quadros 15 a 18 estão relacionadas ao total de 280 artigos e podem sugerir a tendência dos estudos realizados sobre o tema ontologia no campo da CI.

Considerando a associação de palavras-chave a áreas do conhecimento que aproveitam o uso de ontologia para organizar e recuperar informações em seus campos de atuação, destacam-se: *Medicine (medical)*, com 18 ocorrências, que representam 6,4% do total. Nesta área, surge como relevante o periódico *Artificial Intelligence in Medicine*, com 10 artigos indexados. As demais áreas são: *Engineering*, 11 ocorrências (3,9%), *Law*, com 5 ocorrências (1,8%), *Philosophy*, com 4 ocorrências (1,4%), *Chemistry (chemical)*, com 3 ocorrências (1,1%) e *Geographic*, com 2 ocorrências (0,7%).

Os últimos artigos indexados na base LISA, na pesquisa feita no período de maio e de junho de 2006, estão apresentados no quadro abaixo. Todos foram publicados em inglês.

QUADRO 19
Artigos do ano de 2006 indexados na base LISA

AUTOR	TÍTULO DO PERIÓDICO	TÍTULO DO ARTIGO	PALAVRAS-CHAVE
Chen, Y.; Zhou, L.; Zhang, D	Journal of Database Management; v.17, n.1, jan-mar 2006, p.67-84	Ontology-supported Web Service composition: an approach to service oriented knowledge management in corporate services	Knowledge management; Business information; Finance; World Wide Web; Companies; Ontologies
Nilakanta, S.; Miller, L.; Zhu, D.	Journal of Database Management; v. 17, n. 1, jan-mar 2006, p. 85-94	Organizational memory management: technological and research issues	Knowledge management; Organizational; Memory
Qin, J.; Hernandez, N.	Journal of American Society for Information Science and Technology, v. 57, n.2, jan. 2006, p. 280-292	Building interoperable vocabulary and structures for learning objects	Thesauri; Learning materials; Metadata
Jakkilinki, R.; Sharda, N.; Ahmad, I.	Journal of Enterprise Information Management, v.19, n.2, 2006, p. 165-174	MUDPY ontology: a tool for multimedia project planning, design and development	Multimedia; Evaluation; Performance measures; Ontologies; Multimedia design and planning pyramid

Com exceção de Qin, J e Zhang, D, ambos com dois artigos indexados, todos os demais autores aparecem pela primeira vez em 2006. Também é a primeira vez que o periódico *Journal of Enterprise Information* é indexado. Em relação às palavras-chave, os termos/expressões *ontologies*, *knowledge management*, *world wide web* e outros se repetem em relação aos anos anteriores, mas termos/expressões como *business information*, *organizational memory*, *finance* e *performance measures* surgem pela primeira vez em 2006.

4. Conclusões

O estudo realizado sobre a frequência do uso do termo Ontologia na CI procurou apresentar a evolução do volume de publicações sobre o tema a partir da identificação das categorias ano, idioma, autor, periódico e palavra-chave, dentro de uma abordagem quantitativa.

O levantamento bibliométrico foi realizado na base de dados LISA - *Library and Information Abstracts*, considerada a mais importante no espaço informacional da CI. Foram encontrados artigos indexados a partir do ano de 1977 até março de 2006.

Com base nos resultados encontrados, verifica-se uma grande incidência de publicações sobre ontologia – 280 artigos – em 85 periódicos diferentes, sendo 67 em inglês.

Os primeiros artigos indexados sobre o tema foram publicados em 1977, 1980 e 1987. Em 1987 houve uma interrupção e apenas em 1991 voltou a ser publicado outro artigo havendo, a partir daí, pelo menos uma publicação por ano. Em 1995 o número de publicações chega a 21, mas apenas a partir de 2001 observa-se uma tendência crescente de publicações sobre Ontologia. O ano de 2004 desponta como o ano mais produtivo: 54 artigos indexados.

Foram encontradas publicações em 11 idiomas. Entretanto, a preponderância do idioma inglês é marcante: 82,52% das publicações indexadas na base LISA foram em inglês, contra pouco mais de 17% nos demais 10 idiomas encontrados. Excetuando o ano de 1987, em que foi publicado apenas um artigo no idioma italiano, nos demais anos em que existem publicações há artigos em inglês. Só aparecem publicações no idioma russo a partir de 2000.

Publicações nos idiomas africano, português, húngaro, norueguês e japonês não apresentam continuidade na produção. Publicações em russo, chinês e italiano estão indexadas até 2005. Não se pode concluir, entretanto, que autores desses países interromperam sua produção científica, mas pode-se levantar a hipótese de que publicar em inglês é cada vez mais uma exigência da área.

É interessante constatar que, embora a produtividade no idioma inglês seja preponderante tanto em relação ao número de publicações (231 do total de 280) quanto ao número de periódicos indexados, destacam-se um periódico russo e um chinês pela produtividade. Os periódicos com maior número de publicações são: 3 periódicos de idioma inglês: *International Journal of Human-Computer Studies* (50 artigos), *IEEE Intelligent Systems* (33 artigos) e *AI Magazine* (15 artigos), um periódico de idioma russo *Nauchno-Tekhnicheskaya Informatsiya* (15 artigos), um periódico que publica em chinês, *Journal of the China Society for Scientific and Technical Information* (13 artigos), e outro de idioma inglês, *Artificial Intelligence in Medicine*, com 10 artigos publicados.

Da mesma forma, embora o número de autores de língua inglesa seja bastante grande, quando analisada a produtividade dos autores, individualmente, os russos Artem'eva e Kleshchnev –que publicam no único periódico russo indexado – se destacam. Além disso, pode ser constatada a existência um grupo de autores russos que costumam publicar juntos: Artem'eva, Kleshchnev, Knyazeva e Kupnevich. Sugere-se para estudos futuros uma análise do conteúdo desses artigos.

Na análise da categoria palavra-chave, as palavras “*ontologies*” e “*ontology*” apareceram 128 vezes, representando aproximadamente 10% do total de 1252 palavras-chave indexadas. As outras palavras-chave mais citadas se referem a termos relacionadas à rede mundial de computadores, aplicações em computador, inteligência artificial, busca, gestão e representação do conhecimento.

Considerando que a base LISA não traz informações sobre o país de origem do autor, mas apenas sobre o idioma de publicação, estudos devem ser realizados procurando identificar país de origem, parcerias entre autores de origens diferentes, a existência de vínculos dos autores com universidades, temas mais frequentes estudados em cada país e tópicos correlacionados.

Além disso, a quantidade de artigos associada ao grande número e variedade de palavras-chave encontrados neste estudo sugere que alguns questionamentos só poderão ser analisados quando estudos qualitativos aprofundarem os dados quantitativos apresentados. A análise realizada mostra a direção que estudos sobre ontologia vêm

tomando ao longo do período estudado e, da mesma forma, indica tendências a serem exploradas para estudos futuros na busca da compreensão da abrangência das publicações e das grandes áreas do conhecimento relacionadas à ontologia.

A partir desse levantamento bibliométrico, devem ser realizados estudos de análise de conteúdo para que se possa estabelecer, de forma mais aprofundada, uma análise do escopo em que a ontologia se vê inserida, hoje, no campo da Ciência da Informação.

Referências

ALMEIDA, M. B.; BAX, M. P. Uma visão geral sobre ontologias: pesquisa sobre definições, tipos, aplicações, métodos de avaliação e de construção. **Ciência da Informação**, Brasília, v.32, n.3, p.7-20, 2003.

GLOSSÁRIO GERAL DE CIÊNCIA DA INFORMAÇÃO. Disponível em: <<http://www.cid.unb.br/publico/setores/100/123/sistema/m0039015.htm>>. Acesso em 06 jul. 2007.

GRUBER, T. R. **What is an Ontology?** Disponível em: <<http://www.ksl.stanford.edu/kst/what-is-an-ontology.html>>. Acesso em 10 maio 2006.

GUARINO, Nicola. Formal ontology and information systems. In: **PROCEEDINGS OF FOIS**, 1998 6-8 June, Trento, Italy. Proceedings... Amsterdam: IOS Press, 1998. p. 3-15.

HOUAISS, Antonio. **Dicionário Houaiss da língua portuguesa**. Disponível em: <<http://www.houaiss.uol.com.br/busca.jhtm>> para assinantes UOL (www.uol.com.br) mediante login e senha. Acesso em: 10 fev. 2006.

MACIAS-CHAPULA, C.A. O papel da informetria e da cienciometria e sua perspectiva nacional e internacional. **Ciência da Informação**, Brasília, v.27, n.2 p.134–140, maio/ago. 1998.

VITORINO, M.; BRÄSCHER, M. **Aula sobre Ontologia**. Disponível em: <<http://aprender.unb.br>>. Acesso em 23 maio 2006.

IBICT. Bases de Dados com acesso regulamentado por assinatura. Disponível em: <<http://www.ibict.br/secao.php?cat=Biblioteca%20do%20IBICT/fontes/bases>>. Acesso em: 15 jul 2006.

NOTAS:

ⁱ Trabalho apresentado à disciplina “Tesauros, taxonomias e ontologias” do Programa de Pós-Graduação do Departamento de Ciência da Informação, Universidade de Brasília (PPGCINF/UnB)

ⁱⁱ Documentação da Universidade de Brasília em julho de 2006. Orientado pela Prof.^a Dr^a Marisa Bräscher – CID/UnB.

Marta Sianes Oliveira do Nascimento

Mestranda em Ciência da Informação Departamento de Ciência da Informação e Documentação, CID- UnB. martasianes@yahoo.com.br

Geovane Eugênio Oliveira

Especialista em Inteligência Competitiva e Organizacional.

Geraldo Gonçalves Bastos

Mestre em Gestão do Conhecimento e da Tecnologia da Informação, Universidade Católica de Brasília. Consultor na MSA-INFOR SISTEMAS E AUTOMAÇÃO. ggbdino@ig.com.br

Fernanda Passini Moreno

Mestre em Ciência da Informação, Departamento de Ciência da Informação e Documentação, CID- UnB.

Consultora - Ibict

Currículo Lattes: <http://lattes.cnpq.br/7471118579735377> . fernandam@ibict.br

Recebido em: 24/10/2006

Aceito para publicação em: jan. 2007